

ENTERED

RAW SEQUENCE LISTING

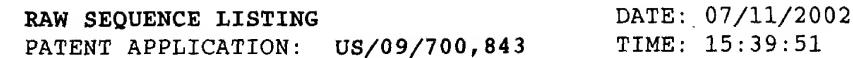
PATENT APPLICATION: US/09/700,843

DATE: 07/11/2002 TIME: 15:39:51

Input Set : A:\Lukacsovich Sequence Listing (Filed Copy).txt

Output Set: N:\CRF3\07112002\I700843.raw

```
3 <110> APPLICANT: LUKACSOVICH, Tamas
              ASZTALOS, Zoltan
              YAMAMOTO, Daisuke
              AWANO, Wakae
      8 <120> TITLE OF INVENTION: A Vector for Gene Trap, and A Method for Gene Trapping by
Using The
              Vector
     11 <130> FILE REFERENCE: 2000-1561A/LC/00653
     13 <140> CURRENT APPLICATION NUMBER: 09/700,843
     14 <141> CURRENT FILING DATE: 2001-02-07
     16 <150> PRIOR APPLICATION NUMBER: PCT/JP99/02683
     17 <151> PRIOR FILING DATE: 1999-05-21
     19 <150> PRIOR APPLICATION NUMBER: JP 10-141952
     20 <151> PRIOR FILING DATE: 1998-05-22
     22 <160> NUMBER OF SEQ ID NOS: 7
     24 <170> SOFTWARE: PatentIn Ver. 2.0
     26 <210> SEQ ID NO: 1
     27 <211> LENGTH: 11206
     28 <212> TYPE: DNA
     29 <213> ORGANISM: Artificial sequence
     31 <220> FEATURE:
W--> 32 <221> NAME/KEY: 3'P sequence
     33 <222> LOCATION: (1)..(237)
     35 <220> FEATURE:
W--> 36 <221> NAME/KEY: synthetic splicing acceptor site and stop/start sequence
     37 <222> LOCATION: (238)..(274)
     39 <220> FEATURE:
W--> 40 <221> NAME/KEY: Gal4 gene (coding region and 3'UTR)
     41 <222> LOCATION: (275)..(3164)
     43 <220> FEATURE:
W--> 44 <221> NAME/KEY: hsp70 terminator
     45 <222> LOCATION: (3165)..(3426)
     47 <220> FEATURE:
W--> 48 <221> NAME/KEY: synthetic junction sequence
     49 <222> LOCATION: 3427-3457
     51 <220> FEATURE:
W--> 52 <221> NAME/KEY: heat shock promoter directed neomycine resistance gene on complementer
              strand
     53
     54 <222> LOCATION: (3458)..(4907)
     56 <220> FEATURE:
W--> 57 <221> NAME/KEY: mini-white gene
     58 <222> LOCATION: (4908)..(8275)
     60 <220> FEATURE:
W--> 61 <221> NAME/KEY: synthetic splicing donor site
```



Input Set : A:\Lukacsovich Sequence Listing (Filed Copy).txt

Output Set: N:\CRF3\07112002\I700843.raw

```
62 <222> LOCATION: (8276)..(8299)
     64 <220> FEATURE:
W--> 65 \langle 221 \rangle NAME/KEY: 5'P sequence
     66 <222> LOCATION: (8300)..(8446)
     68 <220> FEATURE:
W--> 69 <221> NAME/KEY: bacterial part of pCasper3 shuttle vector including complete pUC8
              sequence
W - - > 70
     71 <222> LOCATION: (8447)..(11206)
     73 <220> FEATURE:
W--> 74 <221> NAME/KEY: synthetic DNA
     75 <222> LOCATION: (238)..(274)
              (3427)..(3457)
     76
              (4908)..(4914)(8276)..(8299)
     77
     78 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
     80 <400> SEQUENCE: 1
                                                                                60
     81 catgatgaaa taacataagg tggtcccgtc ggcaagagac atccacttaa cgtatgcttg
     82 caataagtgc gagtgaaagg aatagtattc tgagtgtcgt attgagtctg agtgagacag
                                                                               120
     83 cgatatgatt gttgattaac ccttagcatg tccgtggggt ttgaattaac tcataatatt
                                                                               180
     84 aattagacga aattatttt aaagttttat ttttaataat ttgcgagtac gcaaagctct
                                                                               240
     85 ttctcttaca ggtcgaattg atgtgatgga tccaatgaag ctactgtctt ctatcgaaca
                                                                               300
     86 agcatgcgat atttgccgac ttaaaaaagct caagtgctcc aaagaaaaac cgaagtgcgc
                                                                               360
     87 caagtgtctg aagaacaact gggagtgtcg ctactctccc aaaaccaaaa ggtctccgct
                                                                               420
     88 gactagggca catctgacag aagtggaatc aaggctagaa agactggaac agctatttct
                                                                               480
     89 actgattttt cctcgagaag accttgacat gattttgaaa atggattctt tacaggatat
                                                                               540
                                                                               600
     90 aaaagcattg ttaacaggat tatttgtaca agataatgtg aataaagatg ccgtcacaga
     91 tagattggct tcagtggaga ctgatatgcc tctaacattg agacagcata gaataagtgc
                                                                               660
     92 gacatcatca tcggaagaga gtagtaacaa aggtcaaaga cagttgactg tatcgattga
                                                                               720
     93 ctcggcagct catcatgata actccacaat tccgttggat tttatgccca gggatgctct
                                                                               780
     94 tcatggattt gattggtctg aagaggatga catgtcggat ggcttgccct tcctgaaaac
                                                                               840
     95 ggaccccaac aataatgggt tctttggcga cggttctctc ttatgtattc ttcgatctat
                                                                               900
                                                                               960
     96 tggctttaaa ccggaaaatt acacgaactc taacgttaac aggctcccga ccatgattac
     97 ggatagatac acgttggctt ctagatccac aacatcccgt ttacttcaaa gttatctcaa
                                                                              1020
     98 taattttcac ccctactgcc ctatcgtgca ctcaccgacg ctaatgatgt tgtataataa
                                                                              1080
     99 ccagattgaa atcgcgtcga aggatcaatg gcaaatcctt tttaactgca tattagccat
                                                                              1140
     100 tggagcctgg tgtatagagg gggaatctac tgatatagat gttttttact atcaaaatgc
                                                                               1200
     101 taaatctcat ttgacgagca aggtcttcga gtcaggttcc ataattttgg tgacagccct
                                                                               1260
     102 acatettetg tegegatata cacagtggag geagaaaaca aataetaget ataattttea
                                                                               1320
     103 cagettttee ataagaatgg ceatateatt gggettgaat agggaeetee eetegteett
                                                                               1380
                                                                               1440
     104 cagtgatage ageattetgg aacaaagaeg eegaatttgg tggtetgtet aetettggga
     105 gatccaattg tccctgcttt atggtcgatc catccagctt tctcagaata caatctcctt
                                                                               1500
     106 cccttcttct gtcgacgatg tgcagcgtac cacaacaggt cccaccatat atcatggcat
                                                                               1560
                                                                               1620
     107 cattgaaaca gcaaggctct tacaagtttt cacaaaaatc tatgaactag acaaaacagt
     108 aactgcagaa aaaagtccta tatgtgcaaa aaaatgcttg atgatttgta atgagattga
                                                                               1680
                                                                               1740
     109 ggaggtttcg agacaggcac caaagttttt acaaatggat atttccacca ccgctctaac
     110 caatttgttg aaggaacacc cttggctatc ctttacaaga ttcgaactga agtggaaaca
                                                                               1800
     111 gttgtctctt atcatttatg tattaagaga ttttttcact aattttaccc agaaaaagtc
                                                                               1860
     112 acaactagaa caggatcaaa atgatcatca aagttatgaa gttaaacgat gctccatcat
                                                                               1920
                                                                               1980
     113 gttaagcgat gcagcacaaa gaactgttat gtctgtaagt agctatatgg acaatcataa
     114 tgtcacccca tattttgcct ggaattgttc ttattacttg ttcaatgcag tcctagtacc
                                                                               2040
```

RAW SEQUENCE LISTING DATE: 07/11/2002 PATENT APPLICATION: US/09/700,843 TIME: 15:39:51

Input Set: A:\Lukacsovich Sequence Listing (Filed Copy).txt
Output Set: N:\CRF3\07112002\1700843.raw

				gaatgctgag			2100
				attaaaaaaa			2160
				agaggtatgt			2220
118	gtgtgcaatc	ccattaccgc	atatcagtta	taacaatagt	aatggtagcg	ccattaaaaa	2280
119	tattgtcggt	tctgcaacta	tcgcccaata	ccctactctt	ccggaggaaa	atgtcaacaa	2340
120	tatcagtgtt	aaatatgttt	ctcctggctc	agtagggcct	tcacctgtgc	cattgaaatc	2400
121	aggagcaagt	ttcagtgatc	tagtcaagct	gttatctaac	cgtccaccct	ctcgtaactc	2460
122	tccagtgaca	ataccaagaa	gcacaccttc	gcatcgctca	gtcacgcctt	ttctagggca	2520
123	acagcaacag	ctgcaatcat	tagtgccact	gaccccgtct	gctttgtttg	gtggcgccaa	2580
124	ttttaatcaa	agtgggaata	ttgctgatag	ctcattgtcc	ttcactttca	ctaacagtag	2640
				aaattctcaa			2700
				taatgaaatc			2760
				ttggacggac			2820
128	tggaatcact	acagggatgt	ttaataccac	tacaatggat	gatgtatata	actatctatt	2880
				aaaagagtaa			2940
				cgtgcaccat			3000
				ctcctctaag			3060
				agaattaatg			3120
				ttattcagaa			3180
				ttaagttttt			3240
				ctaatacata			3300
136	gcaacaaatt	gattttagct	atattagcta	cttggttaat	aaatagaata	tatttattta	3360
137	aagataatto	gtttttattg	tcagggagtg	agtttgctta	aaaactcgtt	tagatccact	3420
				aaggagggcg			3480
139	cccacactaa	aggatcatcc	agccggcgtc	ccggaaaacg	attccgaagc	ccaacctttc	3540
				tgatggcagg			3600
				aactcgtcaa			3660
142	cactacaaat	cadaacadc	gataccgtaa	agcacgagga	agcggtcagc	ccattcgccg	3720
143	ccaagetett	cagcaatatc	acgggtagee	aacgctatgt	cctgatagcg	gtccgccaca	3780
144	cccagccggc	cacagtcgat	gaatccagaa	aagcggccat	tttccaccat	gatattcggc	3840
145	aagcaggcat	caccatagat	cacqacqaqa	tectegeegt	cgggcatgcg	cgccttgagc	3900
146	ctggcgaaca	attcaactaa	cacdaacccc	tgatgctctt	cqtccagatc	atcctgatcg	3960
147	acaagaccgg	cttccatccg	agtacgtgct	cgctcgatgc	gatgtttcgc	ttggtggtcg	4020
				agccgccgca			4080
				aggagatcct			4140
				acgtcgagca			4200
				togtoctgca			4260
152	aggtcggtct	tgacaaaaag	aaccgggcgc	ccctgcgctg	acageeggaa	cacqqcqqca	4320
153	tcagaggagg	cgattgtctg	ttgtgcccag	tcatagccga	atageetete	cacccaagcg	4380
				tcaatcatgc			4440
				ttcttgtatt			4500
156	atttcagtag	ttgcagttga	tttacttqqt	tgctggttac	ttttaattga	ttcactttaa	4560
				ttcagctgcg			4620
				tgaattgaat			4680
				cgaacattcg			4740
				agagtgagag			4800
				ataacggcca			4860
				accagtttgt			4920
				tttttggaga			4980
700	codageacyc	acycaayca			,,		

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/700,843

DATE: 07/11/2002
TIME: 15:39:51

Input Set : A:\Lukacsovich Sequence Listing (Filed Copy).txt

Output Set: N:\CRF3\07112002\1700843.raw

164	atttttttt	tattttttac	tgcactggac	atcattgaac	ttatctgatc	agttttaaat	5040
				aggttctttc			5100
166	gacattccac	tcaaagtcag	cgctgtttgc	ctccttctct	gtccacagaa	atatcgccgt	5160
167	ctctttcgcc	gctgcgtccg	ctatctcttt	cgccaccgtt	tgtagcgtta	cctagcgtca	5220
168	atgtccgcct	tcagttgcac	tttgtcagcg	gtttcgtgac	gaagctccaa	gcggtttacg	5280
169	ccatcaatta	aacacaaagt	gctgtgccaa	aactcctctc	gcttcttatt	tttgtttgtt	5340
170	ttttgagtga	ttggggtggt	gattggtttt	gggtgggtaa	gcaggggaaa	gtgtgaaaaa	5400
171	tcccggcaat	gggccaagag	gatcaggagc	tattaattcg	cggaggcagc	aaacacccat	5460
172	ctgccgagca	tctgaacaat	gtgagtagta	catgtgcata	catcttaagt	tcacttgatc	5520
173	tataggaact	gcgattgcaa	catcaaattg	tctgcggcgt	gagaactgcg	acccacaaaa	5580
174	atcccaaacc	gcaatcgcac	aaacaaatag	tgacacgaaa	cagattattc	tggtagctgt	5640
175	gctcgctata	taagacaatt	tttaagatca	tatcatgatc	aagacatcta	aaggcattca	5700
176	ttttcgacta	cattctttt	tacaaaaaat	ataacaacca	gatattttaa	gctgatccta	5760
177	gatgcacaaa	aaataaataa	aagtataaac	ctacttcgta	ggatacttcg	ttttgttcgg	5820
178	ggttagatga	gcataacgct	tgtagttgat	atttgagatc	ccctatcatt	gcagggtgac	5880
179	agcggacgct	tcgcagagct	gcattaacca	gggcttcggg	caggccaaaa	actacggcac	5940
180	gctcctgcca	cccagtccgc	cggaggactc	cggttcaggg	agcggccaac	tagccgagaa	6000
181	cctcacctat	gcctggcaca	atatggacat	ctttggggcg	gtcaatcagc	cgggctccgg	6060
182	atggcggcag	ctggtcaacc	ggacacgcgg	actattctgc	aacgagcgac	acataccggc	6120
183	gcccaggaaa	catttgctca	agaacggtga	gtttctattc	gcagtcggct	gatctgtgtg	6180
184	aaatcttaat	aaagggtcca	attaccaatt	tgaaactcag	tttgcggcgt	ggcctatccg	6240
185	ggcgaacttt	tggccgtgat	gggcagttcc	ggtgccggaa	agacgaccct	gctgaatgcc	6300
186	cttgcctttc	gatcgccgca	gggcatccaa	gtatcgccat	ccgggatgcg	actgctcaat	6360
187	ggccaacctg	tggacgccaa	ggagatgcag	gccaggtgcg	cctatgtcca	gcaggatgac	6420
188	ctctttatcg	gctccctaac	ggccagggaa	cacctgattt	tccaggccat	ggtgcggatg	6480
189	ccacgacatc	tgacctatcg	gcagcgagtg	gcccgcgtgg	atcaggtgat	ccaggagctt	6540
190	tegeteagea	aatgtcagca	cacgatcatc	ggtgtgcccg	gcagggtgaa	aggtctgtcc	6600
191	ggcggagaaa	ggaagcgtct	ggcattcgcc	tccgaggcac	taaccgatcc	gccgcttctg	6660
192	atctgcgatg	agcccacctc	cggactggac	tcatttaccg	cccacagcgt	cgtccaggtg	6720
193	ctgaagaagc	tgtcgcagaa	gggcaagacc	gtcatcctga	ccattcatca	gccgtcttcc	6780
194	gagctgtttg	agctctttga	caagatcctt	ctgatggccg	agggcagggt	agctttcttg	6840
195	ggcactccca	gcgaagccgt	cgacttcttt	tcctagtgag	ttcgatgtgt	ttattaaggg	6900
196	tatctagcat	tacattacat	ctcaactcct	atccagcgtg	ggtgcccagt	gtcctaccaa	6960
197	ctacaatccg	gcggactttt	acgtacaggt	gttggccgtt	gtgcccggac	gggagatcga	7020
198	gtcccgtgat	cggatcgcca	agatatgcga	caattttgct	attagcaaag	tagcccggga	7080
199	tatggagcag	ttgttggcca	ccaaaaattt	ggagaagcca	ctggagcagc	cggagaatgg	7140
200	gtacacctac	aaggccacct	ggttcatgca	gttccgggcg	gtcctgtggc	gatcctggct	7200
201	gtcggtgctc	aaggaaccac	tcctcgtaaa	agtgcgactt	attcagacaa	cggtgagtgg	7260
202	ttccagtgga	aacaaatgat	ataacgctta	caattcttgg	aaacaaattc	gctagatttt	7320
203	agttagaatt	gcctgattcc	acacccttct	tagtttttt	caatgagatg	tatagtttat	7380
204	agttttgcag	aaaataaata	aatttcattt	aactcgcgaa	catgttgaag	atatgaatat	7440
205	taatgagatg	cgagtaacat	tttaatttgc	agatggttgc	catcttgatt	ggcctcatct	7500
206	ttttgggcca	acaactcacg	caagtgggcg	tgatgaatat	caacggagcc	atcttcctct	7560
207	tcctgaccaa	catgaccttt	caaaacgtct	ttgccacgat	aaatgtaagt	cttgtttaga	7620
208	atacatttgc	atattaataa	tttactaact	ttctaatgaa	tcgattcgat	ttaggtgttc	7680
209	acctcagagc	tgccagtttt	tatgagggag	gcccgaagtc	gactttatcg	ctgtgacaca	7740
210	tactttctgg	gcaaaacgat	tgccgaatta	ccgctttttc	tcacagtgcc	actggtcttc	7800
211	acggcgattg	cctatccgat	gatcggactg	cgggccggag	tgctgcactt	cttcaactgc	7860
212	ctggcgctgg	tcactctggt	ggccaatgtg	tcaacgtcct	tcggatatct	aatatcctgc	7920

RAW SEQUENCE LISTING DATE: 07/11/2002 PATENT APPLICATION: US/09/700,843 TIME: 15:39:51

Input Set : A:\Lukacsovich Sequence Listing (Filed Copy).txt

Output Set: N:\CRF3\07112002\I700843.raw

							,
	gccagctcct	_					7980
	ctctttggcg	=					8040
215	tacctctcat	ggttccgtta	cgccaacgag	ggtctgctga	ttaaccaatg	ggcggacgtg	8100
216	gageegggeg	aaattagctg	cacatcgtcg	aacaccacgt	gccccagttc	gggcaaggtc	8160
217	atcctggaga	cgcttaactt	ctccgccgcc	gatctgccgc	tggactacgt	gggtctggcc	8220
218	attctcatcg	tgagcttccg	ggtgctcgca	tatctggctc	taagacttcg	ggcccgacgc	8280
	aaggagtaga						8340
	tccgcacaca						8400
	gtaagcttcg						8460
	agtccagcgg						8520
	tggggcacgt						8580
	ggttaatcag						8640
	ggtatactgg						8700
	ctgacgagca						8760
	aaagatacca						8820
	cgcttaccgg						8880
	cacgctgtag						8940
	aaccccccgt						9000
	cggtaagaca						9060
	ggtatgtagg						9120
	ggacagtatt						9180
	gctcttgatc						9240
	agattacgcg						9300
	acgctcagtg						9360
	tcttcaccta						9420
	agtaaacttg						9480
239	gtctatttcg	ttcatccata	attacctaac	tececateat	gtagataact	acqatacqqq	9540
	agggcttacc						9600
	cagatttatc						9660
	ctttatccgc						9720
	cagttaatag						9780
	cgtttggtat						9840
	ccatgttgtg						9900
	tggccgcagt						9960
	catccgtaag	_					10020
	gtatgcggcg						10080
	gcagaacttt						10140
	tcttaccgct						10200
	catcttttac						10260
	aaaagggaat						10320
	attgaagcat						10380
	aaaataaaca						10440
	aaaccattat						10500
	tegegegttt						10560
	cagettgtet						10620
	ttggcgggtg						10680
	accatatgcg						10740
	gtcgtcgaac						10800
	tacgaaacgt						10860
		55-5-55-55	- 55-5	- 5 5 5 5 5		J . J	

VERIFICATION SUMMARY

PATENT APPLICATION:

US/09/700,843

DATE: 07/11/2002 TIME: 15:39:52

Input Set : A:\Lukacsovich Sequence Listing (Filed Copy).txt

Output Set: N:\CRF3\07112002\1700843.raw

```
L:32 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:36 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:40 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:44 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:48 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:52 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:57 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:61 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:65 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:69 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:70 M:257 W: Feature value mis-spelled or invalid, Describe feature in <223> for SEQ ID#:1
L:74 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
```